HMM solutions

Both the Viterbi and marginal solutions make use of efficient recursive computations. In both cases there is a forward recursion through the Markov chain followed by a backward recursion that traverses the chain in reverse.

It is common to explain these recursive algorithms using matrix multiplication, which reduces the solutions to arithmetic. Presented here are solutions that use simple probability. In particular joint distributions, conditional distributions and Bayes' rule. This approach maintains a strong connection between the two random processes that define the HMM and the solutions given the data.

The following algorithms for implementations of both the Viterbi and marginal solutions are for a HMM that is homogeneous, reversible and irreducible.

Variable definitions

- Q is the set of cytogenetic states modelled by the HMM
- $\mathrm{K} = |\mathrm{Q}|$ is the number of states in Q
- q and q' are states in Q
- T is the number of inputs to the HMM
- $t \in 1, \ldots, \mathrm{T}$ is an index for the inputs
- y_t is the input value at t
- y is the vector $\mathrm{y}_1, \ldots, \mathrm{y}_\mathrm{T}$
- $\mathcal{S}$ is the set of configurations of states
- $\mathrm{K}^\mathrm{T} = |\mathcal{S}|$ is the number of configurations in S
- $l \in 1, \ldots, \mathrm{K}^\mathrm{T}$ is an index for the configurations
- S_l is a configuration in $\mathcal{S}$
- s is a configuration drawn from $\mathcal{S}$
- $\mathrm{p}(\mathrm{s} = \mathrm{S}_l)$ is the probability that a random draw, s, from $\mathcal{S}$, is S_l
- $\mathrm{p}(\mathrm{y}|\mathrm{s} = \mathrm{S}_l)$ is the probability of observing y, given that the random draw, s, from $\mathcal{S}$, is S_l

The Viterbi solution

The Viterbi solution identifies the configuration of hidden states in $\mathcal{S}$ with the greatest posterior probability. That is, $\arg\max_l \mathrm{p}(\mathrm{s} = \mathrm{S}_l|\mathrm{y})$

Forward recursion Both the forward and backward recursion in the Viterbi solution can be motivated by evaluating a single transition as follows.

1. Consider an HMM of length T truncated at t where $1 < t \leq \mathrm{T}$ and let $\mathrm{P}_{\mathrm{t},q}$ be the greatest posterior probability of all configurations of hidden states of length t terminating at state q
2. Likewise, let $\mathrm{P}_{\mathrm{t}-1,q'}$ be greatest posterior probability of all configurations of hidden states of length $t-1$ terminating at state q'
3. Posterior probabilities can be carried forward from $t-1$ to t by accounting for the transition and emission probabilities
$$\mathrm{P}_{t,q} = \max_{q' \in \mathrm{Q}} \left[\mathrm{P}_{t-1,q'}\, \mathrm{p}(\mathrm{s}_t = q | \mathrm{s}_{t-1} = q')\right] \mathrm{p}(\mathrm{y}_t | \mathrm{s}_t = q). \tag{1}$$
4. From Eq (1), it can be determined that the configuration that produces $\mathrm{P}_{t,q}$ has state q' at $t-1$ where
$$q' = \arg\max_{q' \in \mathrm{Q}} \left[\mathrm{P}_{t-1,q'}\, \mathrm{p}(\mathrm{s}_t = q | \mathrm{s}_{t-1} = q')\right]. \tag{2}$$
5. When $t = \mathrm{T}$ the most probable of all configurations terminates at state q where
$$q = \arg\max_{q \in \mathrm{Q}} \left[\mathrm{P}_{\mathrm{T},q}\right]. \tag{3}$$
This configuration is the Viterbi solution.
6. The entire configuration that leads to the Viterbi solution in Eq (3) can be obtained by computing solutions to Eq (1), from $t = 1$ to $t = \mathrm{T}$. Recording solutions to Eq (2) at each step. Then tracing backwards from $t = \mathrm{T}$ to $t = 1$.

Eq (1) is recursive. At each step, two numbers, each with a value less than 1, are multiplied by the result result of the previous step which was also less than 1. Thus, $\mathrm{P}_{t,q}$ quickly takes on a very small value unless the results are normalised in order to retain numerical stability. The following algorithm implements the Viterbi solution with a modification that maintains numerical stability by normalising $\mathrm{P}_{t,q}$ at each recursion.
Step 1: at $t = 1$, for each $q \in \mathrm{Q}$

1. use only the likelihoods
$$\mathrm{A}_q = \mathrm{p}(\mathrm{y}_1 | \mathrm{s}_1 = q)$$
2. rescale for numerical stability
$$\mathrm{a}_{1,q} = \mathrm{A}_q / \sum_{q \in \mathrm{Q}} \mathrm{A}_q$$

Step 2: in the order $t = 2, \ldots, \mathrm{T}$, for each $q \in \mathrm{Q}$

1. transitions and likelihoods
$$\mathrm{A}_q = \max_{q' \in \mathrm{Q}} \left[\mathrm{a}_{t-1,q'} \mathrm{p}(\mathrm{s}_t = q | \mathrm{s}_{t-1} = q')\right] \mathrm{p}(\mathrm{y}_t | \mathrm{s}_t = q)$$
2. rescale to relative probabilities for numerical stability
$$\mathrm{a}_{t,q} = \mathrm{A}_q / \sum_{q \in \mathrm{Q}} \mathrm{A}_q$$

3. extend the most probable path

   $$\mathrm{b}_{t-1,q} = \arg\max_{q' \in \mathrm{Q}} [\mathrm{a}_{t-1,q'} \mathrm{p}(\mathrm{s}_t = q|\mathrm{s}_{t-1} = q')]$$

Step 3: at $t = \mathrm{T}$

1. with forward recursion complete, create a vector $\hat{\mathrm{S}}$ of length T to store the most probable configuration
2. the most probable configuration of hidden states terminates at

   $$\hat{\mathrm{S}}_{\mathrm{T}} = \arg\max_{q \in \mathrm{Q}} [\mathrm{a}_{\mathrm{T},q}]$$

Step 4: in the reverse order $t = \mathrm{T} - 1, \ldots, 1$

1. backward recursion to fill in $\hat{\mathrm{S}}$

   $$q = \hat{\mathrm{S}}_{t+1}$$
   $$\hat{\mathrm{S}}_t = b_{t,q}$$

The Viterbi solution is $\hat{\mathrm{S}}$.

The marginal solution

The marginal solution computes the marginal probability $\mathrm{p}(s_t = q|\mathrm{y})$ for each $t \in 1, \ldots, \mathrm{T}$. The algorithm that produces the solution makes use of the product

$$\mathrm{p}(\mathrm{s}_t = q|\mathrm{y}) \quad \propto \quad \mathrm{p}(\mathrm{s}_t = q|\mathrm{y}_1, \ldots, \mathrm{y}_t)\mathrm{p}(\mathrm{y}_{t+1}, \ldots, \mathrm{y}_{\mathrm{T}}|\mathrm{s}_t = q). \tag{4}$$

Forward recursion is derived from the first term of the product on the RHS of Eq (4). Backward recursion is derived from the second term of the product. The implementation details of forward and backward recursion follow the derivation of Eq (4).

Derivation Why the algorithm works can be understood through its derivation. By Bayes' rule

$$\mathrm{p}(\mathrm{s}_t = q|\mathrm{y})\mathrm{p}(\mathrm{y}) \quad = \quad \mathrm{p}(\mathrm{y}|\mathrm{s}_t = q)\mathrm{p}(\mathrm{s}_t = q) \tag{5}$$

and since $\mathrm{p}(\mathrm{y})$ and $\mathrm{p}(\mathrm{s}_t = q)$ are constants,

$$\mathrm{p}(\mathrm{s}_t = q|\mathrm{y}) \quad \propto \quad \mathrm{p}(\mathrm{y}|\mathrm{s}_t = q). \tag{6}$$

Next partition y to factor the RHS.

$$\begin{aligned} \mathrm{p}(\mathrm{s}_t = q|\mathrm{y}) \quad &\propto \quad \mathrm{p}(\mathrm{y}_1, \ldots, \mathrm{y}_t, \mathrm{y}_{t+1}, \ldots, \mathrm{y}_{\mathrm{T}}|\mathrm{s}_t = q) & (7)\\ &\propto \quad \mathrm{p}(\mathrm{y}_1, \ldots, \mathrm{y}_t|\mathrm{s}_t = q)\mathrm{p}(\mathrm{y}_{t+1}, \ldots, \mathrm{y}_{\mathrm{T}}|\mathrm{s}_t = q) & (8) \end{aligned}$$

Independence of $\mathrm{y}_1, \ldots, \mathrm{y}_t$ and $\mathrm{y}_{t+1}, \ldots, \mathrm{y}_{\mathrm{T}}$ given $\mathrm{s}_t = q$ in Eq (8) is evident in the expansion

$$\mathrm{p}(\mathrm{y}_{t+1}, \ldots, \mathrm{y}_{\mathrm{T}}|\mathrm{s}_t = q) \quad = \quad \sum_{q' \in \mathrm{Q}} \mathrm{p}(\mathrm{y}_{t+1}, \ldots, \mathrm{y}_{\mathrm{T}}|\mathrm{s}_{t+1} = q')\mathrm{p}(\mathrm{s}_{t+1} = q'|\mathrm{s}_t = q) \tag{9}$$

which does not depend on $\mathrm{y}_1, \ldots, \mathrm{y}_t$. Then, as in Eq (6)

$$\mathrm{p}(\mathrm{y}_1, \ldots, \mathrm{y}_t|\mathrm{s}_t = q) \quad \propto \quad \mathrm{p}(\mathrm{s}_t = q|\mathrm{y}_1, \ldots, \mathrm{y}_t) \tag{10}$$

and the RHS of Eq (10) can be substituted into the RHS of Eq (8) which is the RHS of Eq (4).

Forward recursion The forward algorithm solves $\mathrm{p}(\mathrm{s}_t = q|\mathrm{y}_1, \ldots, \mathrm{y}_t)$ in the order $t = 1, \ldots, \mathrm{T}$. The recursion can be obtained from

$$\mathrm{p}(\mathrm{s}_t = q|\mathrm{y}_1, \ldots, \mathrm{y}_t) \quad \propto \quad \mathrm{p}(\mathrm{y}_t|\mathrm{s}_t = q) \sum_{q' \in \mathrm{Q}} \mathrm{p}(\mathrm{s}_{t-1} = q'|\mathrm{y}_1, \ldots, \mathrm{y}_{t-1})\mathrm{p}(\mathrm{s}_t = q|\mathrm{s}_{t-1} = q') \tag{11}$$

Step 1: At $t = 1$, for each $q \in \mathrm{Q}$

1. use only the likelihoods

   $\mathrm{A}_q = \mathrm{p}(\mathrm{y}_1|\mathrm{s}_1 = q)$

2. rescale for numerical stability

   $\mathrm{f}_{1,q} = \mathrm{A}_q / \sum_{q \in \mathrm{Q}} \mathrm{A}_q$

Step 2: For $t = 2, \ldots, \mathrm{T}$, for each $q \in \mathrm{Q}$

1. transitions and likelihoods

   $$\mathrm{A}_q \quad = \quad \mathrm{p}(\mathrm{y}_t|\mathrm{s}_t = q) \sum_{q' \in \mathrm{Q}} f_{t-1,q'} \mathrm{p}(\mathrm{s}_t = q|\mathrm{s}_{t-1} = q')$$

2. rescale for numerical stability

   $$\mathrm{f}_{t,q} \quad = \quad \mathrm{A}_q / \sum_{q \in \mathrm{Q}} \mathrm{A}_q.$$

Backward recursion The computations for backward recursion can be derived by letting $\mathrm{b}_{t,q}$ be proportional to the LHS of Eq (9). That is

$$\mathrm{b}_{t,q} \propto \mathrm{p}(\mathrm{y}_{t+1}, \ldots, \mathrm{y}_\mathrm{T}|\mathrm{s}_t = q). \tag{12}$$

Then using the factorisation in Eq (8)

$$\mathrm{p}(\mathrm{y}_{t+1}, \ldots, \mathrm{y}_\mathrm{T}|\mathrm{s}_{t+1} = q') = \mathrm{p}(\mathrm{y}_{t+1}|\mathrm{s}_{t+1} = q')\mathrm{p}(\mathrm{y}_{t+2}, \ldots, \mathrm{y}_\mathrm{T}|\mathrm{s}_{t+1} = q') \tag{13}$$

which can be substituted into the RHS of Eq (8) to produce the backward recursion,

$$\mathrm{b}_{t,q} \propto \sum_{q' \in \mathrm{Q}} \mathrm{p}(\mathrm{y}_{t+1}|\mathrm{s}_{t+1} = q')\mathrm{b}_{t+1,q'}\mathrm{p}(\mathrm{s}_{t+1} = q'|\mathrm{s} = q). \tag{14}$$

Step 1: at $\mathrm{t} = \mathrm{T}$, for all $q \in \mathrm{Q}$

1. initialise the recursion

   $\mathrm{b}_{\mathrm{T},q} = 1/\mathrm{K}$

Step 2: in the order $t = \mathrm{T} - 1, \ldots, 1$, for all $q \in \mathrm{Q}$

1. backward recursion

   $$\mathrm{B}_q = \sum_{q' \in \mathrm{Q}} \mathrm{p}(\mathrm{y}_{t+1}|\mathrm{s}_{t+1} = q')\mathrm{b}_{t+1,q'}\mathrm{p}(\mathrm{s}_{t+1} = q'|\mathrm{s} = q)$$

2. rescale for numerical stability

   $$\mathrm{b}_{t,q} = \mathrm{B}_q / \sum_{q \in \mathrm{Q}} \mathrm{B}_q$$

Marginal solution The solution to Eq (4) is obtained by multiplying the results of the forward and backward recursions. For $t = 1, \ldots, \mathrm{T}$

$$\mathrm{p}(\mathrm{s}_t = q|\mathrm{y}) = \frac{\mathrm{f}_{t,q}\mathrm{b}_{t,q}}{\sum_{q \in \mathrm{Q}} \mathrm{f}_{t,q}\mathrm{b}_{t,q}}.$$